

## V

## HIV-1/SIVcpz Proteins

Introduction .....	513
Key to annotations in protein alignments .....	513
Gag .....	514
Pol .....	524
Vif .....	542
Vpr .....	546
Tat .....	548
Rev .....	550
Vpu .....	552
Env .....	554
Nef .....	570

The HIV-1/SIVcpz protein alignments presented here contain the same sequences used in the HIV-1/SIVcpz complete genome alignment. Please refer to the tables on pages 127 and 130 for information about these sequences.

### Explanation of Symbols in Alignments

Symbol	Meaning
x	nucleotide missing from codon
#	frameshift, or codon contains N or illegal character
\$	stop codon
- -	domain boundaries
/	protein start point
\	protein end point
V	splice site or exon join
*	cysteine
^^^ [NxS, NxT]	glycosylation site
^*^ [NCS, NCT]	glycosylation site with cysteine
CD4	residue critical for CD4 binding
MHR	major homology region
nls	nuclear localization signal
phos site	phosphorylation site
PKC	protein kinase C binding
Zn-motif	Zinc finger binding motif
M41L etc.	site of common drug resistance mutations in RT

## HIV-1/SIVcpz proteins

[illegible]

HIV-1/SIVcpz  
proteins

### hA/S1-A proteins



HIV-1/SIVcpz  
proteins



		major homology region (HMR)		p24 √ p3	
LQEQIGWNT...NNPFIPVGSIVKEWILGLANKIVIMWVSTSTILLIRQPKPKPRVYDRPYKILAEQA SQPVKAWNTETILLVGMANPDCKTILKALGPAAITLEEMWTACQGVGSPGHZARVIAAMSQ					
E. FR. 83. HX2					
01 AE. CF. 90. 90CF0071					
01 AE. JP. 93. 93JP. NH1					
01 AE. TH. 90. CM240					
01 AE. TH. 93. 93TH9021					
02 AG. CM. 02. 02CM. 1677LE					
02 AG. CM. 97. 97CM. MP807					
02 AG. FR. 91. D3763					
02 AG. NG. - .TENG					
02 AG. SE. 94. SE7812					
03 AS. BY. 09. 98BY10443					
03 AS. RU. 97. XA1163. 2					
03 AS. RU. 98. R098001					
04 CFX. CY. 94. CY032					
04 CFX. GR. 91. 97PCH					
04 CFX. GR. 97. 97PWW					
05 DR. BE. - .V11310					
05 DR. BE. 93. V1961					
05 DR. ES. - .X492					
06 CFX. AU. 96. RFP90					
06 CFX. ML. 96. 95ML127					
06 CFX. ML. 98. 95ML82					
06 CFX. SN. 97. 97S81078					
07 EC. CN. - .CMH1179					
07 EC. CN. 97. 97CN001					
07 EC. CN. 97. CN54					
07 EC. CN. 98. 98CN009					
08 EC. CN. 97. 97CNX. GF					
08 EC. CN. 97. 97CNX. 7F					
08 EC. CN. 97. 97CNX. 9F					
08 EC. CN. 98. 98CN006					
09 CFX. GH. 96. 96GH2911					
09 CFX. SN. 98. 95SN1795					
09 CFX. SN. 98. 95SN788					
09 CFX. US. 98. 98US84057					
10 CD. TZ. 96. 96TZ. BF061					
10 CD. TZ. 96. 96TZ. EF071					
10 CD. TZ. 96. 96TZ. EF110					
11 CFX. CM. 02. 02CM. 4118STN					
11 CFX. CM. 96. 4496					
11 CFX. FR. 99. MP1298					
11 CFX. GR. - .GR17					
12 BE. AR. 97. A32989					
12 BE. BE. 99. 98BE1359					
12 BE. IV. 98. 98BE793					
12 BE. IV. 99. 98BE735					
13 CFX. CM. 02. 02CM. 3226WN					
13 CFX. CM. 96. 1849					
13 CFX. CM. 96. 4164					
14 BG. BS. 00. X475					
14 BG. BS. 00. X477					
14 BG. BS. 00. X623					
14 BG. BS. 99. X397					
15 OLB. TH. 02. 02TH. CUB1331					
15 OLB. TH. 02. 02TH. CUB1322					
15 OLB. TH. 98. 98TH. NUC279					
15 OLB. TH. 99. 99TH. R2399					
16 AZD. FE. 00. K15115009					
16 AZD. FR. 97. 97FR8004					
N. CM. - .YRF106					
N. CM. 95. YRF30					
O. BE. 87. ANT70					
C. CM. - .9MARR009					
C. CM. 91. 91CPS180					
C. SN. 99. 99SME1299					
CPZ. CD. - .CPZANT					
CPZ. CM. 98. CAM5					
CPZ. GA. - .CPZGAB					
CPZ. US. 85. CPZUS					

HIV-1/SIVcpz  
proteins

**zdbA/S-1-A114  
HIV-1/S1Vcpz  
proteins**

[illegible]

HIV-1/SIV/gp120 proteins

HIV-1/SIVcpz  
proteins

		Vpr binding		V6 \
B.FR.83.HAE2	GVETITTP.....QKQEP...IDK...ELV...PUTSRSLFQNPSSQS			501
A1.K2.00.NSR4069	-EGLAS-.....Q...R...Q...A-SI-K...K...LL--			497
A1.K2.94.Q25.17	-E...VS-.....L...Q...K...R...Q...V-K...LL--			499
A1.S2.94.S57253	-E...L...AS-.....P...Q...K...R...Q...S-V-K...LL--			499
A1.X2.97.980201	-E...L...AS-.....P...Q...K...R...Q...S-V-K...LL--			496
A1.X2.97.980201.16	-E...L...AS-.....P...Q...K...R...Q...S-V-K...LL--			496
A1.YG.05.104560.15	-E...L...AS-.....P...Q...K...R...Q...S-V-K...LL--			494
A1.YG.05.104560.37	-E...L...AS-.....P...Q...K...R...Q...S-V-K...LL--			500
A2.CO.97.97GDKS10	-E...L...SS-.....L...N...R...R...PS-T-AI-K...LL--			498
A2.CO.97.97GDKS48	-E...L...SS-.....L...L...ETR...P...R-AI-K...LL--			497
A2.CY.94.94CY017.41	-E...A-.....S.....R.....			499
B.AR.99.ASMAL32	-E...S.....S.....Q...K...R...G...P-D-K...\$			501
B.AU.95.MECQ54	-E...S.....S.....M.....G...SA-K...L			506
B.B0.99.BOL0122	-E...S.....S.....M.....G...SA-K...L			504
B.CN...RL42	-E...S.....S.....M.....G...SA-K...L			500
B.ES.89.S61K15	-E...K-S.....S.....R...R...A-K...L			501
B.GA.88.OVI	-E...S.....S.....T...T...A-K...L			501
B.GB.83.CAM1	-E...S.....S.....K...K...A-K...L			502
B.NL.86.3202A21	-E...S.....S.....K...K...A-K...L			501
B.TH.90.EK132	-E...S.....S.....K...K...A-K...L			502
B.US.83.RE	-E...S.....S.....K...K...A-K...L			501
B.US.90.MEAD160	-E...S.....S.....P...T...K...R...K...S-L-T			497
C.BR.92.92BR025	-E...L...V...P...P...K...R...K...S-L			497
C.EW.00.00EW3891.6	-E...A...S.....P...L...K...R...P-SE-A...SG-L			503
C.EW.96.96EW0502	-E...A...S.....P...L...K...R...P-SE-A...SG-L			505
C.ET.86.ETH2233	-E...A...S.....P...L...K...R...P-SE-A...SG-L			497
C.IL.98.95IS002	-E...A...S.....P...L...K...R...P-SE-A...SG-L			492
C.IN.95.95IN21068	-E...A...S.....P...L...K...R...P-SE-A...SG-L			501
C.IN.99.01IN565.10	-E...A...S.....P...L...K...R...P-SE-A...SG-L			491
C.KR.00.KER2019	-E...A...S.....P...L...K...R...P-SE-A...SG-L			492
C.MM.99.MADU101.3	-E...A...S.....P...L...K...R...P-SE-A...SG-L			493
C.TZ.97.97TZ04	-E...A...S.....P...L...K...R...P-SE-A...SG-L			503
C.TZ.98.98TZ017	-E...A...S.....P...L...K...R...P-SE-A...SG-L			491
C.ZA.01.2134NB	-E...A...S.....P...L...K...R...P-SE-A...SG-L			490
C.ZA.97.97ZA003	-E...A...S.....P...L...K...R...P-SE-A...SG-L			495
C.ZM.96.96ZM651	-E...L...S.....Q...K...K...K...K...L			501
D.CO.83.NUX	-E...L...S.....Q...K...K...K...K...L			498
D.CO.83.NUX	-E...L...S.....Q...K...K...K...K...L			502
D.CO.85.22Z6	-E...V...S.....Q...K...K...K...K...L			497
D.CM.01.01CM.00C9BBY	-E...L...S.....Q...K...K...K...K...L			499
D.KE.01.01KE.NKU3006	-E...L...S.....Q...K...K...K...K...L			503
D.TD.99.NN013	-E...L...S.....Q...K...K...K...K...L			499
D.UG.94.94UG114	-E...L...S.....Q...K...K...K...K...L			496
D.UG.99.99UGA08433	-E...L...S.....Q...K...K...K...K...L			492
D.UG.99.99UGB21875	-E...L...S.....Q...K...K...K...K...L			494
E1.BE.93.VI850	-E...L...S.....P...Q...K...G...P-A-K...\$			495
E1.BE.93.VI850	-E...L...S.....P...Q...K...G...P-A-K...\$			492
E1.BE.93.VI850	-E...L...S.....P...Q...K...G...P-A-K...\$			494
E1.BE.93.VI850	-E...L...S.....P...Q...K...G...P-A-K...\$			495
E2.CM.02.02CM.0016BBY	-E...L...S.....P...Q...K...G...P-A-K...\$			491
E2.CM.95.NP255	-E...L...S.....P...Q...K...G...P-A-K...\$			493
E2.CM.95.NP257	-E...L...S.....P...Q...K...G...P-A-K...\$			497
E2.CM.97.CMS3657	-E...L...S.....P...Q...K...G...P-A-K...\$			491
G.BE.96.DKCB1	-E...L...S.....P...Q...K...G...P-A-K...\$			491
G.CM.01.01CM.4049HAN	-E...L...S.....P...Q...K...G...P-A-K...\$			495
G.FI.93.HB3793.12.1	-E...L...S.....P...Q...K...G...P-A-K...\$			498
G.NG.92.92NG083	-E...L...S.....P...Q...K...G...P-A-K...\$			499
G.SE.93.SEB165	-E...L...S.....P...Q...K...G...P-A-K...\$			496
H.BE.93.VI901	-E...L...S.....P...Q...K...G...P-A-K...\$			499
H.BE.93.VI901	-E...L...S.....P...Q...K...G...P-A-K...\$			500
H.CF.90.90CF056	-E...L...S.....P...Q...K...G...P-A-K...\$			501
J.SE.93.SB7887	-E...L...S.....P...Q...K...G...P-A-K...\$			498
J.SE.94.SB7022	-E...L...S.....P...Q...K...G...P-A-K...\$			498
K.CO.97.EOT511C	-E...L...S.....P...Q...K...G...P-A-K...\$			496
K.CM.96.NP535	-E...L...S.....P...Q...K...G...P-A-K...\$			496

[illegible]

proteins  
zdbA/S1-1-114

## HIV-1/SIVcpz proteins

[illegible]



## HIV-1/SIVcpz proteins

[illegible]

[illegible]

**zdbA/S-1-A114  
HIV-1/S1Vcpz  
proteins**

	D67H_K70R	ASP 110 catalytic site	
B. FR. 83. HX2	SPENPVYTPVFAIKKZLSTNWRKLVDFRELNNKTQDPWEVQLGIPHPAGLKKKSVTVLDVGGAYSPLEDEDPRKVTAFIPISNNHETSGIRYCNVLPQGWKGSALPOSSANTKILEPRQNPQIV		335
01_AE.CF.90.90CF0071		H-SK--R--C--	335
01_AE.JP.93.93JP_NH1		S--S--TK-EM--	333
01_AE.TH.90.CM240		S--S--TK-EM--	340
01_AE.TH.93.93TH9021		S--S--TK-EM--	335
02_AG.CM.02.02CM.1677LE		K--L--A--TK-EM--	337
02_AG.CM.97.97CM.MP807		K--V--A--TK-E--	335
02_AG.FR.91.D766		K--V--A--TK-E--	335
02_AG.NG.-.TENG		K--V--A--TK-E--	335
02_AG.SE.94.SE7612		K--V--A--TK-E--	335
03_AS.BY.05.98BY10443		S-Q--C--E--	335
03_AS.RU.97.XAL163_2		S-Q--C--E--	335
03_AS.RU.98.RU98001		S-Q--C--E--	335
04_CFX.CY.94.CY032		PE--T--PK-E--	334
04_CFX.GR.91.97PVCH		PE--T--V--TK-E--	334
04_CFX.GR.97.97PVWY		PE--T--V--TK-EM--	334
05_DF.BE.-.V11310		I--KE--C--EV--	335
05_DF.BE.93.V1961		I--H--F--C--EM--	335
05_DF.BE.-.X492		I--K--C--E--	335
06_CFX.AU.96.FEP90		I--K--T--I--TK-E--	337
06_CFX.ML.96.95ML127		I--K--T--I--TK-E--	334
06_CFX.ML.96.95ML82		I--K--T--I--TK-E--	334
06_CFX.SN.97.97SE1078		I--K--T--I--TK-E--	339
07_EC.CN.-.CMH1179		I--Y--R--	335
07_EC.CN.97.97CN001		I--Y--R--	335
07_EC.CN.97.CN54		I--Y--L--T--G--	335
07_EC.CN.98.98CN009		I--Y--T--T--G--	335
08_EC.CN.97.97CNX3_GF		K--V--C--	331
08_EC.CN.97.97CNX3_7F		K--V--C--	331
08_EC.CN.97.97CNX3_9F		K--V--C--	331
08_EC.CN.98.98CN006		K--V--C--	331
09_CFX.GH.96.96GH2911		K--V--TK-E--	334
09_CFX.SN.96.96SN1795		K--T--TK-E--	334
09_CFX.SN.96.96SN7805		K--T--TK-E--	335
09_CFX.US.98.98US84057		K--T--TK-E--	339
10_CD.TZ.96.96TZ.BF061		I--Y--Y--QM--	334
10_CD.TZ.96.96TZ.BF071		I--Y--Y--EM--	335
10_CD.TZ.96.96TZ.BF110		I--Y--Y--EM--	332
11_CFX.CM.02.02CM.4118STN		I--KN--T--E--	336
11_CFX.CM.96.4496		I--S--T--TK-ED--	334
11_CFX.FR.99.MP1298		I--S--T--TK-ED--	334
11_CFX.GR.-.GR17		I--S--T--TK-ED--	334
12_BP.AR.97.A32989		K--V--V--	335
12_BP.BE.99.ADMA359		K--V--V--	335
12_BP.UV.99.UR7833		K--V--V--	335
12_BP.UV.99.UR7835		K--V--V--	341
13_CFX.CM.02.02CM.3226WN		KN--V--I--E--	341
13_CFX.CM.96.1849		K--V--T--E--	342
13_CFX.CM.96.4164		V--V--T--E--	341
14_BG.BS.00.X475		S--S--T--TK-E--	334
14_BG.BS.00.X477		S--S--T--TK-E--	334
14_BG.BS.00.X623		S--N--T--TK-E--	334
14_BG.BS.99.X397		S--N--T--TK-EM--	334
15_01B.TH.02.02TH.CUR1331		X--S--T--TK-EM--	340
15_01B.TH.02.02TH.CUR1322		X--S--T--TK-EM--	334
15_01B.TH.96.96TH.NU2879		V--S--C--E--	334
15_01B.TH.99.99TH.R2399		V--S--C--E--	333
16_A2D.FE.00.KTSL15009		S--K--V--	329
16_A2D.FR.97.97FR8004		S--K--V--	335
N.CM.-.YRF106		Q--X--C--K--T--	342
N.CM.95.YRF30		Q--X--C--K--T--	342
O.BE.87.ANTV0		E--G--Q--Q--V--D--ED--EUB--	331
C.CM.-.99MARR009		G--Q--Q--X--T--V--D--EUB--	331
C.CM.91.VU78180		G--Q--Q--V--V--D--EUB--	331
C.SN.99.SEMF1299		G--Q--Q--V--V--D--EUB--	331
CPZ.CD.-.CPZANT		I--TS--Q--V--C--	335
CPZ.CM.98.CAM5		I--TS--Q--V--C--	336
CPZ.GA.-.CPZGAB		I--TS--Q--V--C--	333
CPZ.US.85.CPZUS		I--TS--Q--V--C--	337

HIV-1/SIVcpz  
proteins

## HIV-1/SIVcpz proteins

[illegible]

polymerase motif	T215Y	K219Q
B_FR_83_HAE2	YQYMDIIVGSDLEIGQRTKIRELRQHLLRWLLITPEIKVQKSPFLWMCYELHPDQWTVQPIVPEKDSWTVVDIQXLYGKLNWASQIYPSIKVQGLCKLLRGTKALTEVPLTPEAZELAEANRELL	
01_AE_C9_90_90CP4071	---	---
01_AE_JP_93_93JP_NH1	---	---
01_AE_TH_90_CM240	---	---
01_AE_TH_93_93TH9021	---	---
02_AG_CM_02_02CM_1677LE	---	---
02_AG_CM_97_97CM_MP807	---	---
02_AG_FR_91_D763	---	---
02_AG_NG_--TENG	---	---
02_AG_SE_94_SE7612	---	---
03_AE_BY_09_98BY10443	---	---
03_AE_RU_97_KA1163_2	---	---
03_AE_RU_98_RU98001	---	---
04_CFX_CY_94_CY032	---	---
04_CFX_GR_91_97PVC	---	---
04_CFX_GR_97_97PVC	---	---
05_DF_BE_--V11310	---	---
05_DF_BE_93_V1961	---	---
05_DF_BE_--X492	---	---
06_CFX_AU_96_RP890	---	---
06_CFX_ML_96_95M127	---	---
06_CFX_ML_96_95M182	---	---
06_CFX_SN_97_97S1078	---	---
07_EC_CN_--CMH1179	---	---
07_EC_CN_97_97CN001	---	---
07_EC_CN_97_CN54	---	---
07_EC_CN_98_98CN009	---	---
08_EC_CN_97_97CNX_6F	---	---
08_EC_CN_97_97CNX_7F	---	---
08_EC_CN_97_97CNX_9F	---	---
08_EC_CN_98_98CN006	---	---
09_CFX_GH_96_96GH2911	---	---
09_CFX_SN_96_95SN1795	---	---
09_CFX_SN_96_95SN7805	---	---
09_CFX_US_98_98US84057	---	---
10_CD_TZ_96_96TZ_BF061	---	---
10_CD_TZ_96_96TZ_BF071	---	---
10_CD_TZ_96_96TZ_BF110	---	---
11_CFX_CM_02_02CM_4118STN	---	---
11_CFX_CM_96_4496	---	---
11_CFX_FR_99_MP1298	---	---
11_CFX_GR_--GR17	---	---
12_BP_AR_97_A32989	---	---
12_BP_AR_98_A3MA359	---	---
12_BP_UV_98_UPT833	---	---
12_BP_UV_99_UPT835	---	---
13_CFX_CM_02_02CM_3226WN	---	---
13_CFX_CM_96_1849	---	---
13_CFX_CM_96_4164	---	---
14_BG_B5_00_X475	---	---
14_BG_B5_00_X477	---	---
14_BG_B5_00_X623	---	---
14_BG_B5_99_X397	---	---
15_O1B_TH_02_02TH_CUR1331	---	---
15_O1B_TH_02_02TH_CUR1322	---	---
15_O1B_TH_96_96TH_NU2879	---	---
15_O1B_TH_98_98TH_R2399	---	---
16_A2D_FE_00_KTSLI5009	---	---
16_A2D_FR_97_97FR8004	---	---
N_CM_--YRF106	---	---
N_CM_95_VBF30	---	---
O_BE_87_ANTV0	---	---
C_CM_--99MARS09	---	---
C_CM_91_91P8180	---	---
C_SN_99_99S11399	---	---
CPZ_CD_--CPZANT	---	---
CPZ_CM_98_CAM5	---	---
CPZ_GA_--CPZGAB	---	---
CPZ_US_85_CPZUS	---	---

HIV-1/SIVcpz  
proteins